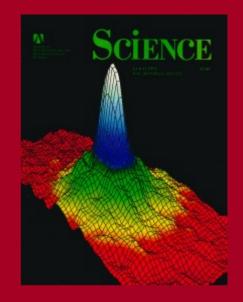
# Outlining a CONSORT statement

Methods – Extension to biobank studies

#### Context

**Epidemiology faces its limits** (Taubes, 1995)





June 22, 2000 Randomized Trials or Observational Tribulations? (Pocock & Elbourne)

A Comparison of Observational Studies and Randomized, Controlled Trials (Benson & Hartz)

Randomized, Controlled Trials, Observational Studies, and the Hierarchy of Research Designs (Concato, Shah & Horwitz)

#### Context

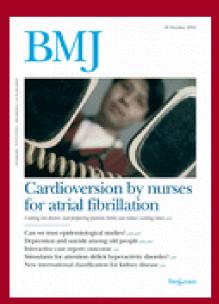
**Beyond randomised versus observational studies** (Concato & Horwitz)

Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? (Lawlor et al.)

When are observational studies as credible as randomised trials? (Vandenbroucke)



22 May, 2004



6 October, 2004

The scandal of poor epidemiological research Reporting guidelines are needed for observational epidemiology (von Elm & Egger)

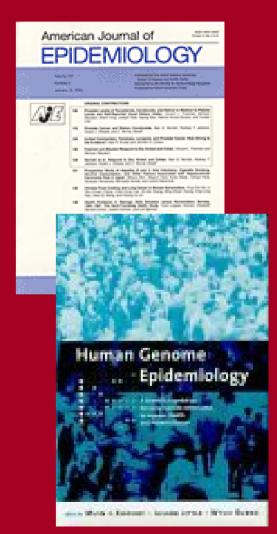
Issues in the reporting of epidemiological studies: a survey of recent practice (Pocock et al.)

# Reporting and Review of Human Genome Epidemiology Studies

- Selection of study subjects
- Analytic validity of genotyping
- Assessment of exposure
- Confounding, including population stratification
- Statistical issues

Reporting, Appraising, and Integrating Data on Genotype Prevalence and Gene-Disease Associations Am J Epidemiol 2002;156:300–10.

Reporting and Review of Human Genome Epidemiology Studies. In: Khoury MJ, Little J, Burke W. (Editors). *Human Genome Epidemiology: A scientific foundation for using genetic information to improve health and prevent disease.* New York, Oxford University Press, 2004, pp. 168-192.

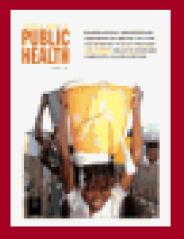


# Checklists for non-randomized evaluations of interventions



Evaluating non-randomised intervention studies (Deeks et al., 2003)

"Although many quality assessment tools exist and have been used for appraising non-randomised studies, most omit key quality domains."



Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement.

(Des Jarlais et al., 2004)

TREND: Transparent Reporting of Evaluations with Nonrandomized Designs

#### Checklists relating to cohort studies



**SIGN 50:** 

A guideline developers' handbook SIGN Publication No. 50, 2001 (updated 2004)

http://www.sign.ac.uk/guidelines/fulltext/50/index.html

Annex C. Critical appraisal: Notes and checklists Methodology Checklist 3: Cohort Studies



Tooth L et al. Quality of Reporting of Observational Longitudinal Research. *Am J Epidemiol* 2005; 161: 280-8.

#### **Participants**

Initial enrolment

**Eligibility criteria** for participants (includes methods of recruitment)

Settings and locations where the data were collected

\*Ethnic group

\*Recruitment from families (e.g. twin pairs, index births and their parents)

\*Nested studies

**Design** – nested case-control, nested case-cohort, nested case-only

<sup>\*</sup>potential issues requiring particular consideration in biobank studies

Types of samples used†

Interventions

Timing of sample collection and analysis†\*

Genotyping

Success rate in extracting DNA†\*

Definition of the genotype(s) investigated; when there are multiple alleles, those tested for should be specified

Genotyping method used (reference; for PCR methods – primer sequences\*, thermocyle profile\*, number of cycles\*)

† Are there differences by study group, e..g. exposure status at enrolment, or in nested studies, between cases and non-cases?

\*Additional information recorded (ideally in web-based methods register)



Quality control measures, including blinding of laboratory staff (to exposure; to outcome in nested studies)†\*#

**Genotyping** contd.

Samples from each group of subjects compared (e.g. cases and non-cases in nested study) included in each batch analyzed\*

† Are there differences by study group, e..g. exposure status at enrolment, or in nested studies, between cases and non-cases?

# See specific heading on blinding (masking)

<sup>\*</sup> Additional information recorded (ideally in web-based methods register)



Methods of assessing exposures documented†

#### **Exposure** assessment

- primary exposures and confounders identified when biobank initiated
- more detailed assessments in nested studies (N.B. recall bias)

Reproducibility and validity of exposure documented

Categories or exposure scale justified

† Are there differences by study group, e..g. exposure status at enrolment, or in nested studies, between cases and non-cases?

#### **Objectives**

Specific objectives and hypotheses.

In biobank study, a major objective (and undertaking!) is establishing the biobank itself.

Some specific objectives and hypotheses formulated *a priori* (for funding agencies; depending on interests of investigators).

Others are likely to be added over time, e.g. as a result of new collaborations. These would be *a priori* hypotheses in the sense that they are not data driven, but may be secondary in the sense that the biobank was not specifically designed to test them.

Objectives contd.

Specific objectives and hypotheses

Potential combination of:

- assessment of large number of genotypes (enabled by high throughput genotyping)
- assessment of large number of exposures assessed at multiple time points
- multiple outcomes

#### **Outcomes**

Clearly defined primary and secondary outcome measures

Compared with RCT, broader range of disease outcomes likely to be assessed in a biobank study (but information about potential complications of intervention, QoL, patient-borne costs unlikely to be sought)

Scale of biobank studies means methods of outcome assessment likely to be less detailed than in RCT, e.g.

- "passive" methods of ascertainment likely to be used, e.g. linkage to cancer registration, hospital discharge data systems, vital records
- self report (positive reports verified by chart abstraction; possibly a sample of negative reports)

Outcomes contd.

When applicable, any methods used to enhance the quality of measurements (eg, multiple observations, training of assessors).

Sample size How sample size was determined

Applies to

- overall design of biobank
- nested studies

Randomization 1 4 1

Confounding

Factors associated with the outcome and exposure under investigation (that are not an intermediate step between exposure and outcome) – data collected and potential confounding assessed in analysis

Alleles associated with the outcome in linkage disequilibrium with the allele under investigation taken into account

#### Randomization

Population stratification:

#### Confounding

•Unaccounted variation in ethnic backgrounds by exposure group when ethnic groups tend to have different exposures and different frequencies of allelic variants

•In nested case-control study, unaccounted variation in ethnic backgrounds of cases and controls, when ethnic groups have different rates of outcome and different frequencies of allelic variants

#### Randomization

Population stratification:

#### Confounding

So far, empirical evidence in populations of European origin suggests magnitude of any bias small (Wacholder et al., 2000; Ardlie et al., 2002; Freedman et al., 2004; Khlat et al., 2004; Wang et al., 2004)

Interpretation of empirical evidence for African American populations mixed (Millikan et al., 2001; Ardlie et al., 2002; Freedman et al., 2004)

Likely to be less of a problem for cohort studies and studies nested within them than for case-control studies.

Blinding (masking)

Whether or not those assessing the outcomes were blinded to exposure status and genotype.

Whether or not those assessing the genotypes

- blinded to exposure status
- in nested study, blinded to outcome

#### Statistical methods

Distinguish clearly a priori hypotheses and hypotheses generated

Statistical methods used to

- Assess associations
- Test for gene-exposure interaction

Methods to take account of

- loss to follow-up
- potential confounding
- missing data

Methods (& justification) for additional analyses, such as subgroup analyses